

RE-DEFINING A LEXICON & LOOKING AT ELECTRONIC DISCOVERY¹

Stuart Allen Garrie, M.D.

I. INTRODUCTION

For any sentence to be intelligible, both speaker and hearer must take one basic idea for granted: that the words in the sentences we speak have the same meanings to both of us. If they do not, we are in a world where nothing is intelligible. The purpose of this paper is to describe a new process for defining the meaning of words beyond the usual historical usage or contextual methods used by current dictionaries and search engines. The new definitional method described in this paper adds precision and clarity to the historical usage definitions. A word is any vocal or written sign,² which connotes a meaning³ Meanings are concepts that we can understand: what we perceive, remember, imagine, reason or conceptualize.⁴

Current dictionaries, the Oxford English Dictionary (“OED”)⁵ or Merriam-Webster Unabridged Dictionary (“MWUD”),⁶ utilize historical word usage to establish the definition of a word. Similarly, the Google⁷ information search engine utilizes historical word usage through page-ranking. The problem with only defining words by historical usage is that after the various usage definitions are given, one still has the conundrum of what the terms in the predicate mean.

For example, the word “hot-dog” in the OED online is defined by historical usage as: “a hot sausage enclosed as a sandwich in a bread roll.”⁸ The word may also mean “expressing delight or strong approval.”⁹ After historical usage has delineated old and new meanings for the word “hot-dog”, the reader must still define the predicate terms in these various usages. The historical usage method does not provide a process to precisely and clearly understand the meaning of the

¹ Stuart Allen Garrie, *Definitional Method to Increase Precision and Clarity of Information (DMTIPCI)*, Non-provisional Patent Application No. 20080167859 (2008).

² George A. Miller, *The Science of Words* 27-31 (1996).

³ Paul Cobley & Litza Jansz, *Introducing Semiotics* (1997).

⁴ Jacques Martain, *An Introduction to Philosophy* 178-238 (1930).

⁵ OXFORD ENGLISH DICTIONARY ONLINE, <http://www.oed.com> (last visited January 2, 2007).

⁶ MERRIAM-WEBSTER UNABRIDGED DICTIONARY CD Ver. 2.5 (2000).

⁷ Sergey Brin & Lawrence Page, *The Anatomy Of a Large-Scale Hypertextual Web Search Engine*, Computer Science Department, Stanford University, Stanford, CA 94305, <http://infolab.stanford.edu/~backrub/google.html> (last visited January 2 2007).

⁸ OXFORD ENGLISH DICTIONARY ONLINE, <http://www.oed.com>, Select “hot-dog” (last visited January 2, 2007).

⁹ *Id.* at Select “hot-dog.”

predicate terms once the predicate terms cannot be further defined by historical usage. This paper proposes a new “*Definitional Method to Increase Precision and Clarity of Information*” (In this essay referred to as “Definitional Method.”).¹⁰

II. THE DEFINITIONAL METHOD

The Definitional Method employs a new approach to improve the precision and clarity of informational processes. Informational processes are dictionaries, digital or paper word classification systems, online data primary-search engines, and languages (the exchange and understanding of words between two or more people). The Definitional Method’s deconstructive procedure refines the meaning of words. The Definitional Method’s constructive procedure establishes a new way of classifying knowledge.

A. DEFINING PREDICATE TERMS BY DECONSTRUCTION

The Definitional Method determines words or meanings that cannot be defined by other words. The Definitional Method deconstructs terms used in definitional predicates until only primary words remain. A word is a primary word if it cannot be defined by other words. The primary words obtained by the Definitional Method are listed in *Appendix A*. The procedure is to exhaustively analyze the words of definitional predicates. In this way, a more precise meaning may be achieved for sentences in general because we have greater clarity in the meanings of the predicate words that define the sentence’s subject.

The Definitional Method’s procedure for understanding the meaning of a primary word is through examples of the primary word.¹¹ A primary word can only be defined (i.e., understood) by examples.¹²

Here is the basic procedure. Take any word and then execute the Definitional Method by focusing on the word-terms in the predicate of a definition from the previously mentioned dictionaries. One then proceeds to define each term in the definition’s predicate; each term in the predicate will have its own definitional predicate. This defining process is iterated until there are no predicate words left that can be defined by other words.

¹⁰ Stuart Allen Garrie, *Definitional Method to Increase Precision and Clarity of Information (DMTIPCI)*, Non-provisional Patent Application No. 20080167859, at 1 (2008).

¹¹ Stuart Allen Garrie, *Definitional Method to Increase Precision and Clarity of Information (DMTIPCI)*, Non-provisional Patent Application No. 20080167859 (2008).

¹² Through application to keyword searching during discovery, it is possible that the keyword list will contain a more precise and contextually relevant set of words with respect to the legal issue. This could in turn produce results that provide value beyond that which might normally be expected.

For example, take the definition of “apple.” Several historically or contextually derived definitions are given in dictionaries. Two definitions of apple as a fruit will be subjected to the Definitional Method’s deconstructive procedures. According to the MWUD, an apple is “the pome fruit of the genus *Malus* . . .” and a “pome fruit” is “a fleshy accessory fruit . . .”¹³ One of the definitions of apple in the OED is “The round firm fleshy fruit of a Rosaceous tree (*Pyrus Malus*) . . .”¹⁴ Applying the Definitional Method,¹⁵ one proceeds to examine the defining words listed in the predicate.¹⁶ One central word in the predicate of the definition for apple is “fruit.”¹⁷ According to one definition in the MWUD, a fruit is “a product of a plant growth useful to man or animal.”¹⁸ One of the OED definitions of a fruit is “the edible product of a plant or tree, consisting of the seed and its envelope.”¹⁹ Here, following the Definitional Method, the properties of a “fruit” are then defined.²⁰ Both the MWUD and the OED definitions of a fruit include the word “product.” According to one of the MWUD definitions, a product is “a substance produced from one or more other substances as a result of chemical changes.” Again following the Definitional Method, one of the MWUD definitions of “produce” is “to give being, form, or shape to.”²¹ According to one of the OED definitions, a “product” is “a thing produced by nature or natural processes.”²² According to one of the OED definitions of “produce,” to produce means “to bring (a thing) into existence from its raw materials or elements, or as the result of a process; to give rise to, bring about, effect, cause, make (an action, condition, etc.).”²³

¹³ MERRIAM-WEBSTER UNABRIDGED DICTIONARY CD Ver. 2.5, Select “apple” (2000).

¹⁴ OXFORD ENGLISH DICTIONARY ONLINE, <http://www.oed.com>, Select “apple” (last visited January 2, 2007).

¹⁵ Stuart Allen Garrie, Definitional Method to Increase Precision and Clarity of Information (DMTIPCI), Non-provisional Patent Application No. 20080167859, at 4 (2008).

¹⁶ MERRIAM-WEBSTER UNABRIDGED DICTIONARY CD Ver. 2.5, Select “apple” (2000).

¹⁷ OXFORD ENGLISH DICTIONARY ONLINE, <http://www.oed.com>, Select “apple” (last visited January 2, 2007).

¹⁸ MERRIAM-WEBSTER UNABRIDGED DICTIONARY CD Ver. 2.5, Select “fruit” (2000).

¹⁹ OXFORD ENGLISH DICTIONARY ONLINE, <http://www.oed.com>, Select “fruit” (last visited January 2, 2007).

²⁰ Stuart Allen Garrie, Definitional Method to Increase Precision and Clarity of Information (DMTIPCI), Non-provisional Patent Application No. 20080167859, at 4 (2008).

²¹ MERRIAM-WEBSTER UNABRIDGED DICTIONARY CD Ver. 2.5, Select “produce” (2000).

²² OXFORD ENGLISH DICTIONARY ONLINE, <http://www.oed.com>, Select “product” (last visited January 2, 2007).

²³ *Id.* at Select “produce.”

What does “to bring into existence” mean? This phrase is similar to the MWUD’s “to give being, form, or shape to.”²⁴ How does this predicate differ from such words as create or make? According to the MWUD “to make” means “to bring (a material thing) into being by forming, shaping, or altering material.”²⁵ According to the OED “make” means “to cause the material or physical existence of by some action.”²⁶ “To cause,” “to bring into,” and “to make” are primary word phrases as these phrases cannot be defined by other predicate words, except by the illogic of tautology—the subject’s word meaning is the same as the predicate’s word meaning. And what is made, caused, or brought into existence? It is “a material thing” or “a physical existence;” i.e., to make, cause, bring into existence a material thing. The MWUD defines “an existence” as “something that exists.”²⁷ OED defines an existence as “the fact or state of existing.”²⁸ These are both examples of tautologous predicates.

The Definitional Method’s procedure has yielded the primary words hidden in the original predicate words defining an “apple.”²⁹ These primary words were obtained after successively defining the original predicate terms used to define “apple.”³⁰

B. DEFINING PRIMARY WORDS USING EXAMPLES

Once the primary words in the predicate for “apple” have been found as illustrated above, the Definitional Method proceeds to the next step of the procedure: define primary words by use of examples. Beginning with the primary words found in the above deconstruction of the predicate terms used to define “apple,” let’s examine one of those primary words—“existence.”³¹ Can examples be given of existence? Of course. Everything we know exists or we would not know it. So we can give many examples of existences, perhaps starting with ourselves; thus we learn the meaning of “being” and what it means “to exist.” To elaborate, “I am Stuart.” or “You are Mary.” These examples give the meaning of

²⁴ MERRIAM-WEBSTER UNABRIDGED DICTIONARY CD Ver. 2.5, Select “produce” (2000).

²⁵ *Id.* at Select “make.”

²⁶ OXFORD ENGLISH DICTIONARY ONLINE, <http://www.oed.com>, Select “make” (last visited January 2, 2007).

²⁷ MERRIAM-WEBSTER UNABRIDGED DICTIONARY CD Ver. 2.5, Select “existence” (2000).

²⁸ OXFORD ENGLISH DICTIONARY ONLINE, <http://www.oed.com>, Select “existence” (last visited January 2, 2007).

²⁹ Stuart Allen Garrie, Definitional Method to Increase Precision and Clarity of Information (DMTIPCI), Non-provisional Patent Application No. 20080167859, at 5 (2008).

³⁰ *Id.* at 5.

³¹ *Id.* at 5.

the primary word “existence.”³² What about examples of “to cause” or “to bring into being?” Here are some examples of causation: “I made that desk.” “You caused the bat to hit my arm, and the bat hitting my arm caused my arm to break.” “You made a wonderful chocolate cake.”³³

What about the word, “material?” What is a material thing? A material thing is anything that possesses extension or dimensions: Me, You, it, that ball, this cake. What is a “non-material” thing? My father is dead, but I have a “memory” of him in my “mind.” The mind consists of perceptions, imaginings (pink elephants that fly), memories, and concepts—in other words, these categories of functions of the brain are called mind and are examples of non-material substances.³⁴

Part of clarifying a definition is to ask what the word is “not.” This was done above with the primary word “material” and its contrary, “non-material.” Regarding the primary word, “existence,” one could have asked, what does it mean to not exist? Aristotle has written that a thing ceases to exist when it loses its essential property, e.g. a person's essential property may be defined as “the ability to think;” and when a person loses that property, that person, who was a person, is not a person anymore.

When does an apple become not an apple? An apple is a fruit and a fruit is an edible product of a plant or tree; therefore, if being edible were to be its essential trait or property, it ceases to exist as an apple when it is not edible. There may or may not be a term for the negative of a word. In the case of a person, a non-person is called a corpse after it loses its essential property; as to the word apple, there no name for its new actuality. A corpse is not a person anymore and an inedible product of a Rosaceous tree is not an apple; both have transformed into another existence.

Dictionaries such as the OED do not use the Definitional Method of deconstructing historically derived predicates to identify their primary words. ³⁵ These dictionaries instead give the various definitions of words by historical usage or context. For example, the phrase “apple of my eye”³⁶ uses the word “apple” in a way that changes its meaning from the meaning deconstructed above. But this usage change in the definition of the word “apple” does not prevent the execution of the Definitional Method on the new predicate terms for this new usage derived definition of “apple.” Historically, new predicates for words are often created by new public usages; and in search engine use, new

³² *Id.* at 5.

³³ *Id.* at 5.

³⁴ Jacques Martain, *An Introduction to Philosophy* 152-153 (1930).

³⁵ Stuart Allen Garrie, Definitional Method to Increase Precision and Clarity of Information (DMTIPCI), Non-provisional Patent Application No. 20080167859, at 5 (2008).

³⁶ OXFORD ENGLISH DICTIONARY ONLINE, <http://www.oed.com>, Select “apple” (last visited January 2, 2007).

predicates are created by the frequency with which users go to a particular site in search of a word's meaning.³⁷

If a new definition for a word arises from historical usage or search page use frequency, this new definition does not change the validity of the Definitional Method nor its ability to more precisely and clearly define its predicate terms.³⁸

The preceding example – applying the Definitional Method to the historically derived definition of the word “apple” – illustrates a new procedure to precisely clarify the meaning of apple.³⁹ The Definitional Method shows how to deconstruct a subject's predicate terms to primary words.⁴⁰ These primary words are predicate terms that can only be defined or understood through the Definitional Method's example procedure.⁴¹

C. THE DEFINITIONAL METHOD'S CONSTRUCTIVE PROCEDURE

The Definitional Method's constructive procedure establishes a new way of classifying knowledge. For example, the primary words material and non-material can be used to classify all things. The constructive procedure allows one to build definitions of words from the primary words' broad categories, see *Appendix B*.⁴²

III. USING THE DEFINITIONAL METHOD PROCEDURES FOR TEXTUAL SEARCHES

Our initial research suggests that by utilizing the Definitional Method in conjunction with keyword searches against large information datasets, one may substantially increase the usefulness of keyword search results.⁴³

³⁷ Sergey Brin & Lawrence Page, *The Anatomy Of a Large-Scale Hypertextual Web Search Engine*, Computer Science Department, Stanford University, Stanford, CA 94305, <http://infolab.stanford.edu/~backrub/google.html> (last visited January 2 2007).

³⁸ Stuart Allen Garrie, Definitional Method to Increase Precision and Clarity of Information (DMTIPCI), Non-provisional Patent Application No. 20080167859, at 6 (2008).

³⁹ *Id.* at 4.

⁴⁰ *Id.* at 4-5.

⁴¹ *Id.* at 5.

⁴² *Id.* at 7-8.

⁴³ Further investigation is needed to support this assertion. The Definitional Method's constructive procedure can be used for knowledge acquisition. An example of a search engine using the Definitional Method's procedures would be the following: If a person were to type into a Definitional Method aware search engine the word “apple,” the search engine would immediately deconstruct “apple” as in the example above. The end result is a potentially more accurate and precise search engine algorithm. Going back to the deconstruction example of the historical usage definition of the word “apple”, one can see the predicate terms “product of a plant growth.” If one types into a typical search engine the word “apple,” no predicate terms regarding “product of a plant growth” are seen under the first ten terms brought up by the search engine; on

Furthermore, there are indications that the Definitional Method can delineate customer interest by breaking down categories of customers and their interests into the primary words that define the customers and their interests (see “apple” example above and footnote 43). Additionally, we have applied the Definitional Method to optimize the precision of search engine results, acting as a calibration tool for online search engines.⁴⁴

The Definitional Method has other possible applications as well. For instance, when the Definitional Method was applied to the legal discovery process it outputs a set of additional search words that substantially increases the potential set of relevant documents.⁴⁵ This increase in accuracy respective to the document set searched has both cost and legal efficiencies.

It would appear that by applying the Definitional Method in an iterative fashion, the accuracy of the collective results will increase as the Definitional Method evolves based on the relevancy success ratio.⁴⁶

the other hand, if one types into the search engine “product of a plant growth,” additional information which may be relevant to an investigator’s or customer’s concerns regarding apples can be found. To date, our research has focused on ad-hoc search and query result comparison and tracking. We are in the process of testing against larger data sets.

⁴⁴ *Id.*

⁴⁵ *Id.*

⁴⁶ Understanding and using keywords in the electronic discovery process is largely based on common sense and a simple understanding of language. The Definitional Method represents a significant evolution in the concept of keyword searching and may substantially increase the accuracy and efficiency of search results in the discovery process. The author believes that this technology has the potential to help lawyers who currently type in full-text keywords and expect to see the most relevant information. This model is flawed because keyword and Boolean searches (such as the keyword search systems provided by Google and Yahoo) result in only 20% of all relevant information being found. The Conference proceedings published during the TREC 2007 Legal Conference presented extensive research validating the 20% finding. See <http://trec-legal.umiacs.umd.edu/>. Our experience to date suggests that a viable solution to this problem is to augment the way in which words are defined by adding the Definitional Method to the search algorithm.

APPENDIX A: LIST OF PRIMARY WORDS

How; What; When; Where; Why; Define or Mean (as in meaning) or Denotes; Trait, Characteristic, Property, Universal or Quality; Is (being), Is as a verb (is and exists) or Being as a noun (A being and existent); Substance (All that is.); Material vs. Non-material; Not; True; False (True being defined as corresponding to what is, which is also one definition of Reality; True may also meaning not contradicting what is self-evident.); Necessary; Sufficient; Complete, Whole; Limited (spatially or by other traits); Unlimited (spatially or by other traits); Potential; Actual; Change (in time or existence); Parts Of Speech(Subject; Object; Nouns; Pronouns; Verbs; Prepositions; Copulas); Join vs. Separate; Ought vs. Is (Prescriptive vs. Descriptive); Include vs. Exclude; To Think; I; Me; Of (property related to a thing); To (direction or possession “to itself”); Related, Connected; Opposite; Same, Identical; Class, Set; Elements (in Class or Set); Whole vs. Part(s); All; Every; Some; Any; Many; Unique or Particular (The) vs. General or Universal (A); Order vs. Disorder; Unique or Specific; To do; To make; To use; And; Or; If; If-then; Only; Possession: Belong to, Related to, Part of; First; Beginning; Response; Mirror image(Left- Right); Increase; Decrease; More; Less; More or Greater than; Less than; In vs. Out; On or Over or Above vs. Under or Below; Quantity or Amount; One; Add; Subtract; Unity; Continuous vs. Discontinuous; Absence; Equal; Point; Plane; Clockwise vs. Counter clockwise; After vs. Before; Then; Now vs. Later; Contact; Feelings (Love vs. Hate; Sad vs. Happy; Anger vs. Calm); Aware vs. Unaware; Conscious vs. Unconscious, Will (Free Will).

APPENDIX B. CATEGORIZATION OF PRIMARY WORDS

Most primary words can be grouped under two categories of primary words: (1) Space-Time and (2) Existence.

SPACE-TIME

Space is composed two primary terms:

(1) Position (Location or Direction) and Boundary. Position has a subset of primary terms such as “on,” “under,” “over,” “toward,” “above,” “below,” “away,” “close,” “up,” “down,” “near,” “far,” “left-right” (mirror image).

(2) Extension or Length. Length has a subset of primary terms such as “point,” or “line.” An example of extension is “I am 6 feet tall”.

The concept of “boundaries” or “limited by” can be further subdivided:

(1) Limited by space—by a point, line, or direction. For example, an object could be above a point or below a line; or a space could be thought of as between lines;

(2) Limited by amount—that is by number. For example, there could be 10 oranges to eat or 3 oranges to eat;

(3) Limited by predicate terms. For example, G-d cannot be given any predicate terms as then G-d becomes limited.

Position, location, direction plus extension and length are predicate terms that give meaning to the word space. Units of extension or length are arbitrary designations. For example, 1cm is an arbitrary amount of length “from” here “to” there on a ruler. The following are examples of boundary or position words: The book was on the table and the cat was under the table.

Time is composed of three primary terms: *Past, Present, and Future*. Arbitrary units of Past, Present, and Future are chosen to measure time. For example, one tick of a clock is arbitrarily defined as 1 minute; thus 1 minute by convention or common agreement becomes the fundamental unit of telling time.

Basic space-time concepts can be used to develop fields of knowledge, e.g. the sciences. One of the science fields, physics, uses the terms force, velocity, speed and acceleration. *Force* is a primary word. An example of force as a primary word is weight; specifically, “I weigh 180lbs”. Units of weight are arbitrary. Certain sized pieces of metal that are arbitrarily defined as 1 pound or 1 kilogram are kept by governments as keepers of the standard.⁴⁷

⁴⁷ OXFORD ENGLISH DICTIONARY ONLINE, <http://www.oed.com>, Select “pound” and “kilogram” (last visited January 2, 2007).

EXISTENCE ⁴⁸

Primary words under this second major category follow: is, subject, object, I, me, potential, actual, limits or without limits (limits meaning predicate terms), substance (all that is) and material vs. non-material. An example of “infinite” defined under existence refers to a subject not having any limiting characteristics (predicate terms such as properties, traits, terms, universals). When “is” is used in a sentence, it is referred to as a copula, connecting the subject to the predicate terms.

⁴⁸ Stuart Allen Garrie, *Definitional Method to Increase Precision and Clarity of Information (DMTIPCI)*, Non-provisional Patent Application No. 20080167859, at 7-8 (2008)-- There are five acts of existence: a. Processes: The process for quantity is increasing or decreasing. The process for movement is change in position. The process for cause is becoming or bringing into existence—to be labeled a "cause" the subject bringing something into existence must be not only necessary but sufficient for what is brought into being. b. Linkage primary words: is, set, subset, equal, under, over, if-then. c. Energy can be potential or actual. d. Change can be in time, amount, or existence. Modes: Existence in one’s mind, existence outside one’s mind.